We claim:

1. A method of classifying a human colorectal tumor comprising the steps of:

obtaining an unknown sample derived from a human colorectal tumor;

determining the gene expression level of each of at least five informative genes in the unknown sample, wherein informative genes are selected from the group consisting of:

a first group of genes that are expressed at higher levels in Dukes' C tumors than in Dukes' B tumors, wherein the first group consists of H2BFH, H2BFJ, LENEP, SPON1, PC, PADI2, SPAM1, PHKA1, NKX2H, RAB7, SPN, BAIAP3, DVL2, GRB14, SRCAP, GTF3C1, RUNX1, SYNGR3, PPL, NPIP, BPAG1, ARF6, PNMA1, UPLC1, BAI3, FLJ22596, FLJ10251, KIAA1827, CLIPR-59, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp586C1923, cDNA DKFZp434K1126, DNCH1, DRPLA, ADPRTL2, HSPCA, YWHAZ, AP2A2, GRWD, BICD1, MARK3, FLJ10482, FLJ11588, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416, and

a second group of genes that are expressed at lower levels in Dukes' C tumors than in Dukes' B tumors, wherein the second group consists of FUT4 , NDUFB7, ASMT, CASP5, PSMF1, PTGER3, FLB0708, RPS21, KIAA0734, FZD3, EIF1A, NEBL, SRI, KCNJ6, ANKRD5, SSX2, AF098968, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, cDNA DKFZp434E0528, PTPRA, BS69, CREG, PIGL, ILVBL, PANK2, ATP5J, CGI-01, and UBE2N;

comparing the gene expression level of each of the at least five informative genes in the unknown sample to the average gene expression level of that gene in a plurality of reference samples that are Dukes' B stage to determine if each of the at least five informative genes is expressed at higher or lower levels in the unknown sample relative to the plurality of reference samples; and,

classifying the unknown sample as Dukes' C stage if each of the informative genes selected from the first group is expressed at higher levels in the unknown sample than the average expression of that gene in the plurality of reference samples and if each of the informative genes selected from the second group is expressed at lower levels in the unknown sample than the average expression of that gene in the plurality of reference samples.

2. The method of claim 1 wherein at least twenty informative genes are analyzed at each step.

3. The method of claim 2 wherein the at least twenty informative genes are NKX2H, H2BFJ, PADI2, SPAM1, PHKA1, KIAA1827, PSMF1, ASMT, ARF6, PTGER3, BPAG1, SRI, GTF3C1, KIAA0903, FLJ22596, DT1P1B, DKFZp586C1923, UPLC1, ANKRD5, and KIAA0992.

4. The method of claim 1 wherein the gene expression level of each of the at least five informative genes is determined by hybridization to an array of nucleic acid probes.

5. The method of claim 4 wherein the array of nucleic acid probes comprises SEQ ID NOs 102-321.

6. A method of identifying compounds that inhibit or promote metastasis of a colorectal tumor to the regional lymph nodes comprising:

obtaining a first sample of cells from a colorectal tumor before treatment;

treating the tumor with the compound;

obtaining a second sample of cells from the tumor after treatment;

determining the expression level of at least five informative genes in the first and second samples, wherein the at least five informative genes are selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, FUT4, PC, NDUFB7, PADI2, SPAM1, PHKA1, NKX2H, RAB7, ASMT, CASP5, PSMF1, SPN, PTGER3, FLB0708, RPS21, BAIAP3, KIAA0734, DVL2, GRB14, FZD3, SRCAP, GTF3C1, RUNX1, EIF1A, NEBL, SYNGR3, PPL, NPIP, BPAG1, SRI, KCNJ6, ARF6, PNMA1, UPLC1, BAI3, SSX2, FLJ22596, FLJ10251, AF098968, KIAA1827, CLIPR-59, ANKRD5, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp434E0528, cDNA DKFZp586C1923, cDNA DKFZp434K1126, PIGL, ILVBL, DNCH1, DRPLA, UBE2N, ADPRTL2, CGI-01, HSPCA, PANK2, ATP5J, YWHAZ, AP2A2, GRWD, BICD1, PTPRA, MARK3, BS69, CREG, FLJ10482, FLJ11588, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416;

determining the direction of change between the expression level of each of the at least five informative gene in the first and second samples;

identifying the compound as a promoter of metastasis if the direction of change for each of the at least five informative genes is:

up and the gene is selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, PC, PADI2, SPAM1, PHKA1, NKX2H, RAB7, SPN, BAIAP3, DVL2, GRB14, SRCAP, GTF3C1, RUNX1, SYNGR3, PPL, NPIP, BPAG1, ARF6, PNMA1, UPLC1, BAI3, FLJ22596, FLJ10251, KIAA1827, CLIPR-59, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp586C1923, cDNA DKFZp434K1126, DNCH1, DRPLA, ADPRTL2, HSPCA, YWHAZ, AP2A2, GRWD, BICD1, MARK3, FLJ10482, FLJ11588, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416; or

down and the gene is selected from the group consisting of FUT4 , NDUFB7, ASMT, CASP5, PSMF1, PTGER3, FLB0708, RPS21, KIAA0734, FZD3, EIF1A, NEBL, SRI, KCNJ6, ANKRD5, SSX2, AF098968, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, cDNA DKFZp434E0528, PTPRA, BS69, CREG, PIGL, ILVBL, PANK2, ATP5J, CGI-01, and UBE2N;or,

identifying the compound as an inhibitor of metastasis if the direction of change for each of the at least five informative genes is:

down and the gene is selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, PC, PADI2, SPAM1, PHKA1, NKX2H, RAB7, SPN, BAIAP3, DVL2, GRB14, SRCAP, GTF3C1, RUNX1, SYNGR3, PPL, NPIP, BPAG1, ARF6, PNMA1, UPLC1, BAI3, FLJ22596, FLJ10251, KIAA1827, CLIPR-59, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp586C1923, cDNA DKFZp434K1126, DNCH1, DRPLA, ADPRTL2, HSPCA, YWHAZ, AP2A2, GRWD, BICD1, MARK3, FLJ10482, FLJ11588, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416; or

up and the gene is selected from the group consisting of FUT4 , NDUFB7, ASMT, CASP5, PSMF1, PTGER3, FLB0708, RPS21, KIAA0734, FZD3, EIF1A, NEBL, SRI, KCNJ6, ANKRD5, SSX2, AF098968, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, cDNA DKFZp434E0528, PTPRA, BS69, CREG, PIGL, ILVBL, PANK2, ATP5J, CGI-01, and UBE2N.

7. The method of claim 6 wherein at least twenty informative genes are analyzed at each step.

8. The method of claim 7 wherein the at least twenty informative genes are NKX2H, H2BFJ, PADI2, SPAM1, PHKA1, KIAA1827, PSMF1, ASMT, ARF6, PTGER3, BPAG1, SRI, GTF3C1, KIAA0903, FLJ22596, DT1P1B, DKFZp586C1923, UPLC1, ANKRD5, and KIAA0992..

9. The method of claim 6 wherein gene expression level is determined using an array of nucleic acid probes.

10. A method of predicting the efficacy of a compound for treating a colorectal tumor comprising the steps of:

obtaining a first sample of cells derived from a colorectal tumor before treatment with the compound and a second sample of cells derived from the same tumor after treatment;

determining a gene expression level for at least five informative genes in the first and second sample, wherein informative genes are selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, FUT4, PC, NDUFB7, PADI2, SPAM1, PHKA1, NKX2H, RAB7, ASMT, CASP5, PSMF1, SPN, PTGER3, FLB0708, RPS21, BAIAP3, KIAA0734, DVL2, GRB14, FZD3, SRCAP, GTF3C1, RUNX1, EIF1A, NEBL, SYNGR3, PPL, NPIP, BPAG1, SRI, KCNJ6, ARF6, PNMA1, UPLC1, BAI3, SSX2, FLJ22596, FLJ10251, AF098968, KIAA1827, CLIPR-59, ANKRD5, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp434E0528, cDNA DKFZp586C1923, cDNA DKFZp434K1126, PIGL, ILVBL, DNCH1, DRPLA, UBE2N, ADPRTL2, CGI-01, HSPCA, PANK2, ATP5J, YWHAZ, AP2A2, GRWD, BICD1, PTPRA, MARK3, BS69, CREG, FLJ10482, FLJ11588, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416;

comparing the gene expression level of each of the at least five informative gene in the first and second samples to determine a direction of change in expression from the first to the second sample for each gene, wherein the direction is either up or down;

comparing the direction of change for each gene to the direction of change for that gene from a sample of a first known stage to a sample of a second known stage, wherein the first stage is more advanced than the second stage; and

identifying the compound as effective if the direction of change for each of the at least five informative genes is the same.

11. The method of claim 10 wherein gene expression levels are determined using an array of nucleic acid probes.

12. The method of claim 10 wherein at least twenty informative genes are analyzed at each step.

13. The method of claim 12 wherein the at least twenty informative genes are NKX2H, H2BFJ, PADI2, SPAM1, PHKA1, KIAA1827, PSMF1, ASMT, ARF6, PTGER3, BPAG1, SRI, GTF3C1, KIAA0903, FLJ22596, DT1P1B, DKFZp586C1923, UPLC1, ANKRD5, and KIAA0992.

14. A method of screening to identify test substances which induce or repress expression of genes which are induced or repressed in a colorectal tumor sample that has metastasized compared to a colorectal tumor sample that has not metastasized, comprising:

contacting a colorectal tumor cell with a test substance;

monitoring expression of a transcript or its translation product wherein the transcript is from a gene selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, FUT4, PC, NDUFB7, PADI2, SPAM1, PHKA1, NKX2H, RAB7, ASMT, CASP5, PSMF1, SPN, PTGER3, FLB0708, RPS21, BAIAP3, KIAA0734, DVL2, GRB14, FZD3, SRCAP, GTF3C1, RUNX1, EIF1A, NEBL, SYNGR3, PPL, NPIP, BPAG1, SRI, KCNJ6, ARF6, PNMA1, UPLC1, BAI3, SSX2, FLJ22596, FLJ10251, AF098968, KIAA1827, CLIPR-59, ANKRD5, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp434E0528, cDNA DKFZp586C1923, cDNA DKFZp434K1126, PIGL, ILVBL, DNCH1, DRPLA, UBE2N, ADPRTL2, CGI-01, HSPCA, PANK2, ATP5J, YWHAZ, AP2A2, GRWD, BICD1, PTPRA, MARK3, BS69, CREG,

FLJ10482, FLJ11588, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416;.

15. A method of distinguishing between a Dukes' B and a Dukes' C stage tumor comprising: monitoring the expression level of any five or more genes selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, FUT4, PC, NDUFB7, PADI2, SPAM1, PHKA1, NKX2H, RAB7, ASMT, CASP5, PSMF1, SPN, PTGER3, FLB0708, RPS21, BAIAP3, KIAA0734, DVL2, GRB14, FZD3, SRCAP, GTF3C1, RUNX1, EIF1A, NEBL, SYNGR3, PPL, NPIP, BPAG1, SRI, KCNJ6, ARF6, PNMA1, UPLC1, BAI3, SSX2, FLJ22596, FLJ10251, AF098968, KIAA1827, CLIPR-59, ANKRD5, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp434E0528, cDNA DKFZp586C1923, cDNA DKFZp434K1126, PIGL, ILVBL, DNCH1, DRPLA, UBE2N, ADPRTL2, CGI-01, HSPCA, PANK2, ATP5J, YWHAZ, AP2A2, GRWD, BICD1, PTPRA, MARK3, BS69, CREG, FLJ10482, FLJ11588, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416; and comparing the expression levels to a database of expression levels of the five or more genes in Dukes' B and Dukes' C stage tumors.

16. The method of claim 15 wherein the expression level of at least twenty genes are monitored and compared.

17. The method of claim 16 wherein the at least twenty genes are NKX2H, H2BFJ, PADI2, SPAM1, PHKA1, KIAA1827, PSMF1, ASMT, ARF6, PTGER3, BPAG1, SRI, GTF3C1, KIAA0903, FLJ22596, DT1P1B, DKFZp586C1923, UPLC1, ANKRD5, and KIAA0992..

18. A method of classifying a colorectal tumor sample as being positive for regional lymph node metastases comprising:

isolating a nucleic acid sample from the colorectal tumor sample;

determining the expression level of at least five informative genes in the sample, wherein informative genes are selected from the group consisting of H2BFH, H2BFJ, LENEP, SPON1, FUT4, PC, NDUFB7, PADI2, SPAM1, PHKA1, NKX2H, RAB7, ASMT, CASP5, PSMF1, SPN, PTGER3, FLB0708, RPS21, BAIAP3, KIAA0734, DVL2, GRB14, FZD3, SRCAP,

GTF3C1, RUNX1, EIF1A, NEBL, SYNGR3, PPL, NPIP, BPAG1, SRI, KCNJ6, ARF6, PNMA1, UPLC1, BAI3, SSX2, FLJ22596, FLJ10251, AF098968, KIAA1827, CLIPR-59, ANKRD5, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp434E0528, cDNA DKFZp586C1923, cDNA DKFZp434K1126, PIGL, ILVBL, DNCH1, DRPLA, UBE2N, ADPRTL2, CGI-01, HSPCA, PANK2, ATP5J, YWHAZ, AP2A2, GRWD, BICD1, PTPRA, MARK3, BS69, CREG, FLJ10482, FLJ11588, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416;

comparing the expression level of the at least five informative genes to the expression level of the same informative genes in at least one reference colorectal tumor sample that is negative for regional lymph node metastases;

determining if the one or more informative gene in the colorectal tumor sample is expressed at a higher level or a lower level relative to the reference; and

classifying the sample as having regional lymph node metastases if the direction of change in the expression of the at least five informative genes in the unknown sample relative to the reference is the same as the direction of change for that gene in Tables 2 and 3.

19. A kit for classifying a colorectal tumor sample as being positive for regional lymph node metastases comprising:

an array of probes wherein each probe is perfectly complementary to a single gene selected from the group of genes consisting of H2BFH, H2BFJ, LENEP, SPON1, FUT4, PC, NDUFB7, PADI2, SPAM1, PHKA1, NKX2H, RAB7, ASMT, CASP5, PSMF1, SPN, PTGER3, FLB0708, RPS21, BAIAP3, KIAA0734, DVL2, GRB14, FZD3, SRCAP, GTF3C1, RUNX1, EIF1A, NEBL, SYNGR3, PPL, NPIP, BPAG1, SRI, KCNJ6, ARF6, PNMA1, UPLC1, BAI3, SSX2, FLJ22596, FLJ10251, AF098968, KIAA1827, CLIPR-59, ANKRD5, KIAA0903, KIAA0992, FLB4237, clone HRC00953, cDNA DKFZp434E0528, cDNA DKFZp586C1923, cDNA DKFZp434K1126, PIGL, ILVBL, DNCH1, DRPLA, UBE2N, ADPRTL2, CGI-01, HSPCA, PANK2, ATP5J, YWHAZ, AP2A2, GRWD, BICD1, PTPRA, MARK3, BS69, CREG, FLJ10482, FLJ11588, DJ971N18.2, HSU79252, SP329, FLJ20420, DT1P1B, ELOVL5, EUROIMAGE 29222, and DKFZP564A2416; and wherein the array comprises at least 2 probes for each gene in the group of genes;

a computer-readable medium having computer-executable instructions for performing a method comprising: comparing the gene expression level of at least three informative genes in an experimental sample to the expression level of the corresponding gene in a plurality of samples of known Dukes' stage wherein the Dukes' stage is selected from Dukes' B and Dukes' C and determining if the experimental sample is Dukes' B or Dukes' C stage; and

a computer-readable medium having a plurality of gene expression level values for each of the genes in the group of genes in a plurality of colorectal tumor samples of known Dukes' stage.